A Virtual Reality Video System for Deep Ocean Remotely Operated Vehicles

Eric J. Martin*, Benjamin Erwin*, Kakani Katija*, Amy Phung†, Everardo Gonzalez†, Susan Von Thun*, Heidi Cullen*, and Steven H.D. Haddock*

* Monterey Bay Aquarium Research Institute
Moss Landing, California
Email: emartin@mbari.org

† Olin College of Engineering
Needham, Massachusetts

***Abstract*—Telepresence has always been a core aspect in the design and use of Remotely Operated Vehicles (ROVs) for deep-sea ocean exploration since they have replaced manned submersibles as the preeminent technology for deep sea research. Operators and users of these platforms require significant training to be able to navigate the environment through a multitude of sensors and two-dimensional video presentations. In this arrangement, conventional video cameras do not provide spatial context of study objects since they are flattened in presentation, limiting scientific observations. Display of these many video feeds also necessitates construction of large and costly control rooms, which in turn requires space on constrained vessels. Using multiple two-dimensional perspectives force operators to develop spatial relationships to perform manipulations. A new camera system has been developed at the Monterey Bay Aquarium Research Institute through a collaborative development with Blue Ring Engineering. It is deployed from an ROV, and designed specifically for generating video for virtual reality head-mounted displays. The resultant system facilitates collection of three-dimensional, 180° field-of-view stereo video delivered in real-time to the control room for science class ROVs. The imaging system, ROVVR, is a single 4000 m depth-rated housing with two 4K resolution cameras with fisheye lenses spaced at near-human interpupillary distance. As part of this effort, we have developed a Virtual Control Room (VCR), which presents the three-dimensional VR feed, with two-dimensional standard ROV video feeds as foreground panels, along with contextual data using augmented reality. Together, our software and hardware solutions can be used to improve vehicle piloting, reduce the time it takes for subsea manipulations, in-situ coarse size estimations, and improve portability of deployed systems.**

I. Introduction

Since 1989, the Monterey Bay Aquarium Research Institute has conducted science operations in the deep sea, employing remotely operated vehicles as a primary exploration and discovery platform. These teleoperated robots help scientists explore the deep ocean more safely and efficiently than their predecessor, the manned submersible [1]. Scientific ROVs are required to provide the highest fidelity measurements across all sensors. No sensor engages the attention of operators (or pilots) and researchers more than the video systems on these platforms. These cameras are typically broadcast-quality sensor modules fit to lenses that provide a wide view ($> 80^\circ$) with zoom capabilities [2]. ROV video has been a consistent source of materials used in outreach and ocean engagement as well. These videos have inspired stewardship, conservation, and a general appreciation for the deep ocean [3].

Fig. 1. The ROV control room aboard the R/V *Western Flyer*.

A common problem for all ROVs is that these wide angle cameras still do not provide enough context for the pilot to be truly aware of the entire subsea environment. The current solution to this problem is to judiciously use more cameras, placed where space allows on the vehicle, and aimed at providing video for regions of interest. “Fisheye” ultra-wide cameras can be used to improve subsea vehicle context, however they have proven difficult to interpret when presented on two-dimensional video monitors due to distortion effects. Through regular calibration, these effects can be somewhat removed, but complications in manufacturing remain [4]. Multiple cameras as a solution forces a user to divert their attention across many monitors in the teleoperation space, known as the ROV control room (Figure 1). The user must also have the ability to piece together these disparate video sources in such a way as to provide spatial awareness for the subsea vehicle.

Here we present a system (ROVVR) that combines emerging Virtual Reality (VR) tools with science-class remotely operated vehicles (ROVs). In Section I-B, we present limitations of conventional ROV video displays, which motivate the development of ROVVR. A discussion of the design and integration of the system is presented in Section II. Operational benefits from using VR are discussed in Section II-A. Section III will present use cases and improvements to both human-in-the-loop control and real-time science tools, as well as the impact VR will have on ocean engagement and outreach.

A. Previous Efforts

Since ROVs are fundamentally a teleoperated platform, efforts to bring underwater video systems into head mounted displays have been attempted before. While these efforts have showed great promise [5], they have not gained community-wide adoption due to earlier technology limitations in ROV video quality and head mounted display sophistication. More recently, many efforts have taken place on small ROVs, leveraging the monoscopic video commonly available [6].

A common difficulty in the use of head mounted displays is the development of motion-sickness in operators. This fatigue can be associated with multiple factors: (1) degree of motion capture, (2) the interpupillary distance (IPD) of the display being well matched to the user, and (3) cameras whose inter camera distance (ICD) best approximates human IPD [7], [8]. These findings were thoroughly explored during the first boom of head mounted displays (HMDs) in the 1990's. Since the reinvigoration of virtual reality in the domestic entertainment market, and the wide availability of consumer grade head mounted displays, it is clear that earlier technological limitations (e.g., advent of compact, high quality video sensors and video displays) have been reduced significantly. Improvements to graphics processing units are also critical to reducing latency, which can also contribute to VR fatigue.

B. Conventional ROV Video Challenges

In order to access deep marine waters of 4000 m, MBARI uses ROVs that have changing tool-sleds that are dependent on the dominant science needs of principal investigators. The VR efforts at MBARI have initially involved ROV *Doc Ricketts* (Figure 2) as well as ROV *MiniROV*, and can be operated independent of variable tool-sleds for system flexibility [9].

More generally, ROVs require propulsion in order to navigate their environments, and thrusters provide control through 5 degrees of freedom. Most ROVs are designed to maintain neutral buoyancy to allow for maximum authority in propulsion. Strategic placement in flotation and ballast help create stable vehicles with strong righting moments. Another common aspect of science-class ROVs are manipulator arms, to aid in sample collection, site investigations, and other physical tasks [10]. Particularly relevant to deep sea robotics, no light penetrates below 200-meters (aphotic zone) [11], so all illumination for navigation and observations must be provided by the ROV itself [12]. Power and telemetry are provided for vehicles at these depths through a single tether with conductors and fiber-optic cables, respectively.

The fiber-optics cabling inside these "tethers" provide bidirectional communication and allow for COTS telemetry solutions to transceive a multitude of electrical information both

Fig. 2. The ROV *Doc Ricketts*.

down to the ROV (control) and up to the ship (sensors, video). Common sensors on modern ROVs that are necessary for successful operations include inertial navigation sensors (INS), compasses, pressure sensors (depth), and tracking beacons (for subsea geolocation). Acoustic sensors provide scalar measurements such as altitude, and also three-dimensional target imaging, operating similarly to shipboard radar technologies.

Finally, many lower-quality cameras are placed around higher-definition ones to provide a complete visual picture for an operator to monitor. ROVs often are designed for all interactions to take place on the front of the platform. As such, most light sources are oriented in the same direction. To help extend the utility of the limited high-definition cameras, they are often placed on pan-and-tilt units to help extend the field of view. These actuators commonly do not respond quickly and smoothly, and are often controlled separately from the yaw of the vehicle, making target tracking difficult. Also, the light distribution in the front of the ROV is often set statically, which helps provide the most even light pattern for targets in the center axis of the ROV. Movement of the focal point of these primary cameras often pushes subjects outside of optimal illumination. Many of these shortcomings can be addressed by utilizing modern applications of virtual reality.

II. A Deep Sea 180° 3D VR Camera

In 2019, we collaborated with Blue Ring Engineering, a virtual reality company specializing in underwater 3D and VR imaging technologies, to design a system that will generate a live VR feed to a HMD for use in ROV operations. The system also provides functionality for onboard recording video, providing content for future science and outreach products. Our ROV-compatible VR system uses a two-camera setup, each equipped with a fisheye lens capable of providing 180 hemispherical degrees of coverage (Figure 3). The VR system is mounted to the upper frame of ROV *Doc Ricketts* (Figure 6). This position optimally places the system to provide a complete view of the front of the vehicle and the illuminated scene. The video feeds from each camera are synced and then converted to a fiber optic signal that is transmitted on specific wavelengths of light through ROV *Doc Ricketts's* fiber optic system. Power and ethernet communications are provided by the science interface module on the ROV. In the ROV control room, the signal from each camera is converted back into an HD-SDI signal and sent to a capture and playback card (Blackmagic). This card saves the full-resolution 4K video feed from each camera for later processing and archival purposes, and streams a real-time 2K video feed to a high-performance computer. The virtual control room system described in Section II-A combines these feeds in real-time and displays them to create a 3D environment for a VR headset. This workflow is depicted in Figure 5.

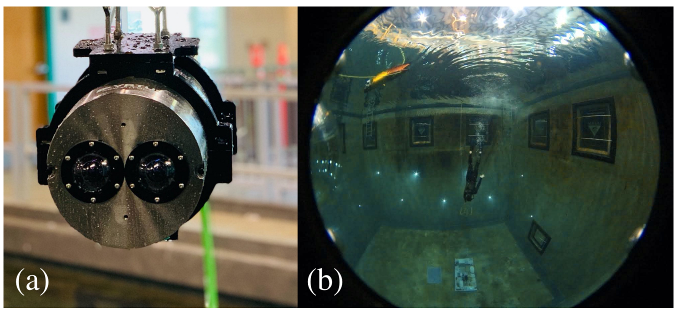

Fig. 3. The ROVVR system (a) and a sample image (b) taken in the MBARI Test Tank facility.

A. The Virtual Control Room

Currently, modern ROV control rooms require significant amounts of dedicated ship space for multiple computers and monitors, where each computer-monitor pair displays a single camera feed or graphic visualization of telemetry data. Using a VR headset condenses the hardware necessary for ROV piloting into a single desktop and VR headset. A VR-based control room enables pilots to leverage the wide field view footage from the fisheye cameras more effectively since they can view the footage without the distortion that results from projecting it onto a 2D screen. Therefore, we developed a VR-based control room (VCR) in Unity that overlays real-time ROV telemetry data on an immersive 1:1-scale stereoscopic view from the camera system described in Section II. In the existing control room, ROV telemetry data (e.g. heading, depth, altitude, etc.) is transferred from the ROV to the ship through a fiber optic cable. Once onboard the ship, the data is converted into Lightweight Communications and Marshalling (LCM) [13] packets that are distributed across the ship's local network. The control room desktop computer receives these packets and overlays them on the stereoscopic view created by the stereo cameras, which makes the data viewable from the VR headset. This data pipeline is illustrated in Figure 5.

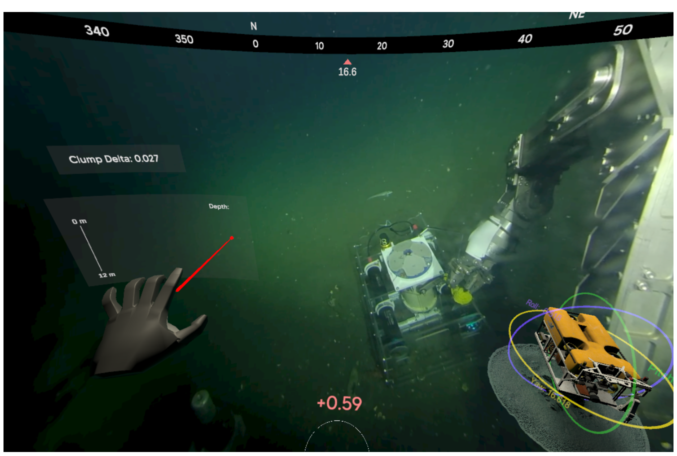

Fig. 4. View of the virtual control room from within the VR headset. Heading is displayed across the top, attitude is displayed in the lower right-hand corner, the tether turns count is displayed on the bottom, and the ROV depth and distance from its tether weight is displayed on the left. Hand tracking enables these foreground displays to be easily rearranged.

In order for the video feeds to be displayed properly in VR, two important pre-processing steps are needed for each camera feed: undistortion and reprojection. The undistortion step uses a third-degree polynomial to account for imperfections in the fisheye lens and create an image that closely resembles the ideal fisheye projection. The reprojection step takes the adjusted image and reprojects it to a hemispherical projection using Bourke's process [14], which is needed to properly display it in the headset. The reprojected image from the right camera is used to create the virtual environment for the right eye, and similarly the reprojected image from the left camera is used for the left eye. The cameras are mounted as close to the average person's IPD as physical constraints from the pressure housing would allow; however, since each person's IPD is slightly different, an x and y offset is added to each image to mitigate the "double-vision" effect that results from this difference. The result is a 3D, 1:1-scale stereoscopic view in VR that provides pilots with depth perception, and gives them an immersed vehicle perspective.

There are also numerous affordances associated with integrating the telemetry data overlays into the VR environment. In the VCR implementation, we used the Leap Motion VR Developer kit to enable pilots to rearrange the virtual telemetry data overlays with their hands. This adds flexibility to how pilots display data, improving upon the fixed nature of the monitors in modern control rooms. The virtualized environment also lets pilots save individual display preferences, allowing them to quickly load their preferred configuration when starting their piloting shifts. While existing control rooms are limited to the 2D displays for telemetry data, the VR control room can utilize 3D features to create more intuitive interfaces. In our implementation, we utilize this to display heading as a 3D ring and display attitude by changing the orientation of a miniature 3D model of the ROV (Figure 4).

The current iteration of the VCR can allow for significant cost savings by reducing the amount of equipment, ship space, and cognitive load posed by traditional ROV control rooms. Additionally, it provides anyone in the virtual control room with increased situational awareness and the ability to tailor their own work environment based on their professional needs.

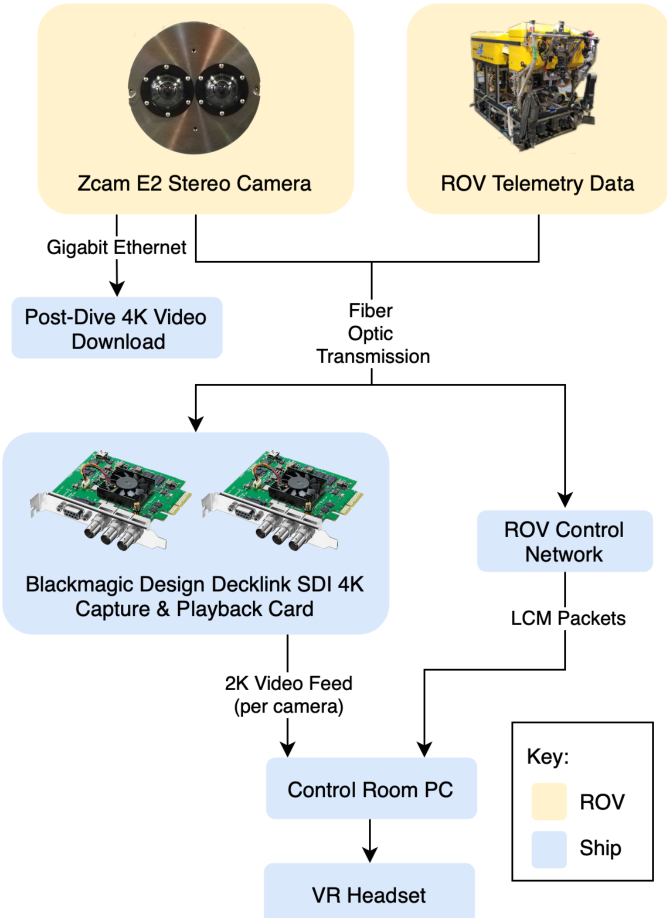

Fig. 5. Data pipeline for the virtual reality system.

Utilizing the 3D panoramic camera system as a backdrop to telemetry overlays allows for the virtualization of data placed within a real setting (Mixed Reality). Future iterations can make improvements by overlaying more complex information such as distant georeferenced waypoints, historical data, and topographical maps on top of the panoramic backdrop.

The current application is limited to a single user in VR at one time. A companion application has been developed that allows other users outside of the HMD to view and collaborate with the VR user. A simple pointer tool exists as well that can be used in a way to "paint" targets of interest for the user in virtual reality.

III. Potential Applications

This generation of subsea hemispherical stereo video can be leveraged for new scientific insights and outreach, and further work in this area will be a primary focus now that the system is fully integrated onto the ROV. Our multidisciplinary team of engineers and scientists have interests in a number of areas, including exploring improvements to vehicle control, sampling, and in-situ video analysis products.

In addition, a number of other efforts have begun to leverage the newest generation of head-mounted displays to provide new haptic control modes for robots, and ROVs in particular [15], [16]. These projects have focused on the possibility of building better human-interface devices.

A. Manipulation Accuracy

Robotic manipulator arms (Figure 6) are the primary tool for an ROV pilot in almost all use cases. Stereo vision systems have been adapted onto ROVs in the past since stereo vision gives the user the perspectives needed to bring a manipulator arm to its target in 3D space [17], [18]. ROV pilots are skilled at leveraging the many perspectives from different cameras to develop sufficient spatial awareness for manipulations. However, these disparate cameras often do not capture all useful views, and operating errors are often made. Some errors merely expend valuable in-water time, but other errors can result in visual loss or damage of delicate samples and even damage to the vehicle due to collisions or entanglement.

B. Decision Support

Aside from time lost to visual scanning of multiple monitors, ROV dives often experience delays due to a lack of spatial awareness of the environment surrounding the vehicle. Due to transmission losses in seawater, ROV lights illuminate roughly 10 meters in front of an ROV. This lack of visibility means that operators must rely on geolocation from acoustic tracking to understand where the vehicle is in space. This new video system, combined with mixed reality tools, should enable the creation of an artificial landscape in the virtual space. This virtual space can present the vehicle relative to a priori obtained seafloor topography, real-time path histories, sample locations, and historical data. Future efforts will involve evaluating whether virtual spaces can support more efficient decision-making during subsea ROV operations.

C. Portability

Space aboard expeditionary vessels is limited in a number of ways. Current ROV control rooms require multiple monitors to display information to operate the vehicle. Beyond operators, principal investigators require these views to make decisions critical to guiding ROV operations. However, with virtual reality, the "top-side" equipment needed to display video, software, and control tools can be much reduced and aided by augmented reality data overlays of the stitched footage. This solution could easily increase in scale and number of participants if footage is simultaneously brought into multiple head-mounted displays.

D. Science Uses

1) Advance a detailed quantitative view of marine communities: Present image-based quantification methods can give an estimate of organisms per time, and a gross approximation of number per volume. To understand species interactions such as predator-prey and mate-finding, this requires knowing the 3D spatial distribution of animals, including nearest-neighbor distances. Calibrated VR footage will enable quantification of a number of features in video, including average distances

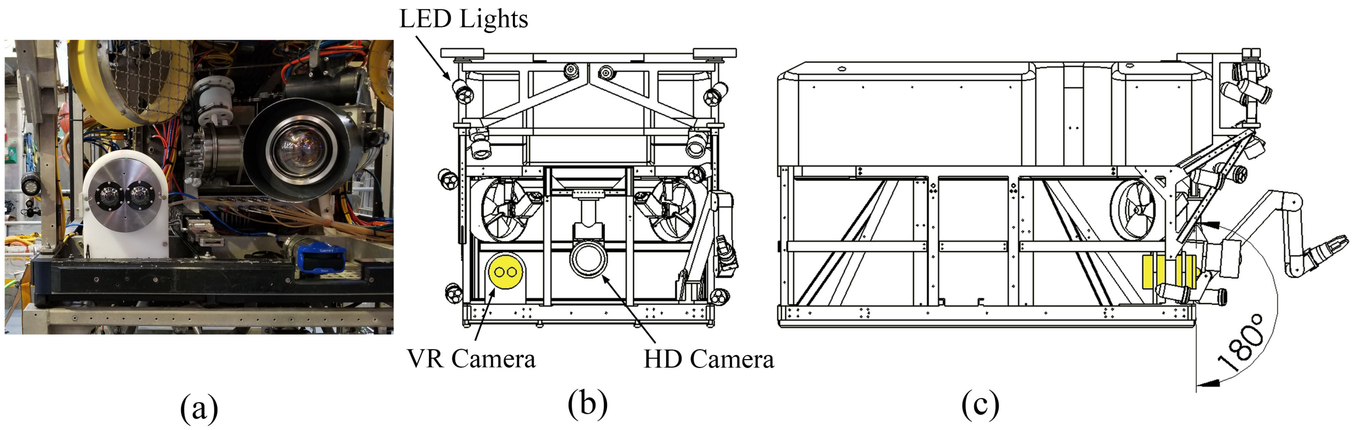

Fig. 6. (a) The ROV *Doc Ricketts* with the VR camera system. (b) Front and (c) side views of the mount are shown with annotations denoting the relative positions of the lighting and HD camera, and the new cameras' field of view.

from organisms in the field of view. This information has the potential to revolutionize our understanding of how animals are interacting in the deep sea, where visual signaling and more covert modes of communication are under-explored.

2) Enhance understanding of the fluctuations and fate of forage species: Within the captured 3D panoramic scene, our camera system will allow scientists to view a recorded deep water scene with contextual information. For example, our extended goals include augmented reality (AR) software tools and a calibrated camera system. This will allow scientists to measure and potentially overlay that information on items in the scene, enabling quantification of biomass as well as raw abundance estimates.

3) Map the seafloor in support of biological and geological investigations: Increased support of our work in creating Photogrammetric 3D models of deep water features using existing ROV cameras, will allow us to repeatedly generate detailed information of both biology, and geology (Figure 7)[1].

4) Devise informatics tools for understanding the ocean and its inhabitants: Post processed 3D video from our system can be overlayed with graphics and text. Visual elements can even become interactive and historical data can be layered. This will allow content viewers the ability to look at any item in the scene, interact with it, and learn more information through "pop-up" text windows or cued audio. Future efforts will involve merging MBARI's AI-driven computer vision elements to the 3D live view for autonomous identification and sizing of elements that enter the scene [19].

E. Outreach

With the ability to walk around and interact freely within captured media and 3D data, VR has the potential to be the ultimate tool to view the ocean and its processes. This new system will present content that allows MBARI and its collaborators to create immersive experiences that emplace users in this unique and mostly unreachable environment. VR is being widely explored for its use in a variety of education and training experiments. Case studies of its use in classroom environments are readily available, and some even with overlap for issues such as climate change, ocean health, [20] and exploration and discovery [21]. Better education encourages stewardship of the oceans, a societal need expressed thoroughly by the oceanographic community [22], and this VR camera system can provide unprecedented access to mostly inaccessible places in the deep ocean.

[1]Model can be viewed online at: https://skfb.ly/6KTNJ

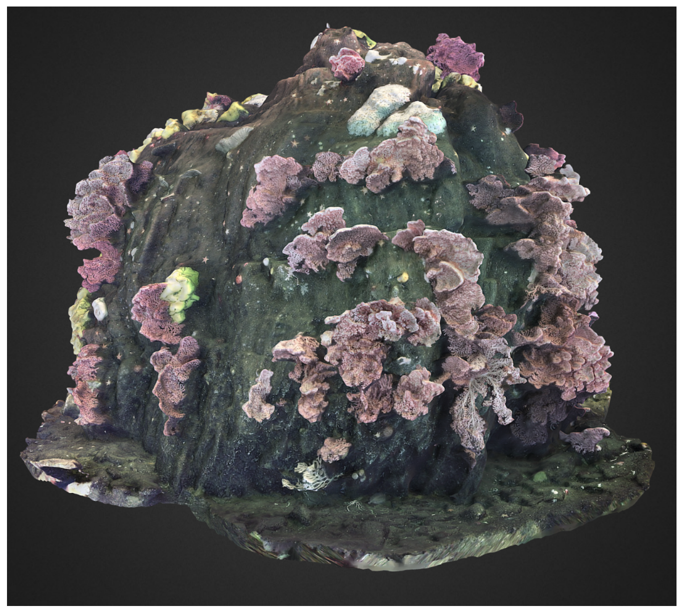

Fig. 7. Photogrammetric reconstruction of a large boulder in the Sur Ridge segment of Monterey Bay.

IV. Conclusions

Virtual reality applications have re-entered the public space and continue to become a more mainstream technology, as many HMD vendors and developers build a more established ecosystem of products. Here we have presented how VR will address a number of needs for exploring the deep sea using remotely operated vehicles, thereby addressing gaps in ROV users' ability to interact fully with the environment in the deep-ocean. The ROVVR camera system will be evaluated for many of the applications presented here.

The system as-built does possess some limitations. Stretching a 4K resolution camera across such a large field of view leads to distortions around the extents of the view. Also, targets often seem pixelated as a result of the display limits of the head mounted displays. A common point of feedback from early trials of the system is that user-perceived resolution does not provide enough detail on small targets. These issues can be addressed in part through upgrades to both camera modules and HMDs as next generation devices become available.

In our early use of this system, we have observed promising improvements across many aspects of ROV use. ROV pilots have reported a sense of better emplacement in the marine environment, and a higher sense of situational awareness. Those that have performed subsea manipulations have felt the stereo enabled VR video is an improvement over the conventional approaches. We have begun efforts with researchers in human-robot-interface science to quantify those improvements.

Collaborating scientists have also expressed enthusiasm for the virtual reality system. Scientists have commented on the utility of a much better sense of scale for the underwater spaces that they study. The ability to observe a wider field of view is of great value when studying complex ecosystems.

We do not expect this system to be a panacea for all the challenges to ROV-based ocean exploration, but these motivations were chosen because they best leverage the system that is practical to develop at this time. The system could be scaled, thereby enabling more immersive, 360°, three-dimensional experiences. This system, as described, enables the MBARI Virtual Reality team the tools needed to evaluate how to best aim future developments, and explore deep waters in new ways.

Acknowledgments

Support for the ROVVR project at MBARI is provided by the David and Lucile Packard Foundation. We would like to thank the pilots of the ROV *Doc Ricketts* for their continued support. We would also like to thank fellow Olin SCOPE team members Erika Lu, Nathan Shuster, and Cameron Wierzbanowski, and their advisor Dr. Lynn Andrea Stein for their contributions to the VR-based control room. Support for the Olin SCOPE team was provided by La Fondation Dassault Systèmes.

References

[1] A. G. Sharp and L. A. Shumaker, *DSRV ALVIN : a review of accomplishments*. Woods Hole Oceanographic Institution, 1977.

[2] B. H. Robison, K. R. Reisenbichler, and R. E. Sherlock, "The co-evolution of midwater research and ROV technology at MBARI," *Oceanography*, vol. 30, no. 4, pp. 26–37, 2017.

[3] J. P. Delgado, H. K. Van Tilburg, B. G. Terrell, D. Marx, C. Marzin, S. Gittings, W. Kiene, V. Grussing, and P. Orlando, "How NOAA's Office of National Marine Sanctuaries engages the public in the ocean through the science and management of maritime heritage," *Aquatic Conservation: Marine and Freshwater Ecosystems*, vol. 26, no. S2, pp. 200–212, 2016. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/aqc.2643

[4] C. Hughes, P. Denny, E. Jones, and M. Glavin, "Accuracy of fish-eye lens models," *Applied Optics*, vol. 49, no. 17, pp. 3338–3347, 2010.

[5] B. P. H. Iii, C. Stoker, M. Sims, D. Rasmussen, T. W. Fong, J. Steele, D. Barch, E. Miles, and E. N. M. I. T, "The Application of Telepresence and Virtual Reality to Subsea Exploration," *Sensors (Peterborough, NH)*, no. March, 1993.

[6] M. Candeloro, E. Valle, M. R. Miyazaki, R. Skjetne, M. Ludvigsen, and A. J. Sørensen, "HMD as a new tool for telepresence in underwater operations and closed-loop control of ROVs," in *OCEANS 2015 - MTS/IEEE Washington*, oct 2015, pp. 1–8.

[7] S. Best, "Perceptual and oculomotor implications of interpupillary distance settings on a head-mounted virtual display," *IEEE Proceedings of the National Aerospace and Electronics Conference*, vol. 1, pp. 429–434, 1996.

[8] J. P. Wann, S. Rushton, and M. Mon-Williams, "Natural problems for stereoscopic depth perception in virtual environments," *Vision Research*, vol. 35, no. 19, pp. 2731–2736, oct 1995. [Online]. Available: https://www.sciencedirect.com/science/article/pii/004269899500018U?via{\%}3Dihub

[9] Monterey Bay Aquarium Research Institute, "Remotely operated vehicles — MBARI," p. 1, 2021. [Online]. Available: https://www.mbari.org/at-sea/vehicles/remotely-operated-vehicles/

[10] S. Sivčev, J. Coleman, E. Omerdić, G. Dooly, and D. Toal, "Underwater manipulators: A review," *Ocean Engineering*, vol. 163, pp. 431–450, sep 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0029801818310308

[11] A. P. Trujillo, *Essentials of oceanography*, 9th ed. Upper Saddle River, N.J.: Pearson Education, 2008.

[12] R. Ballard, D. Yoerger, W. Stewart, and A. Bowen, "Argo/jason a remotely operated survey and sampling system for full-ocean depth," in *OCEANS 91 Proceedings*, vol. 1, 1991, pp. 71–75.

[13] A. S. Huang, E. Olson, and D. C. Moore, "Lcm: Lightweight communications and marshalling," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 4057–4062.

[14] P. Bourke, "Omni-directional Stereoscopic Fisheye Images for Immersive Hemispherical Dome Environments," in *Annual International Conferences on Computer Games, Multimedia and Allied Technology*, 2009.

[15] H. Martins, I. Oakley, and R. Ventura, "Design and evaluation of a head-mounted display for immersive 3D teleoperation of field robots," *Robotica*, vol. 33, no. 10, pp. 2166–2185, 2015.

[16] T. E. Solstad, I. Skjetne, Roger, and I. Candeloro, Mauro, "Improved user-experience for control of ROVs," Master's thesis, Norwegian University of Science and Technology, 2016.

[17] Pan-Mook Lee, Bong-Hwan Jeon, Seok-Won Hong, Yong-Kon Lim, Chong-Moo Lee, Jong-Won Park, and Chang-Min Lee, "System design of an ROV with manipulators and adaptive control of it," in *Proceedings of the 2000 International Symposium on Underwater Technology (Cat. No.00EX418)*, may 2000, pp. 431–436.

[18] S. Ishibashi, "The stereo vision system for an underwater vehicle," in *OCEANS 2009-EUROPE*, may 2009, pp. 1–6.

[19] O. Boulais, B. Woodward, B. Schlining, L. Lundsten, K. Barnard, K. C. Bell, and K. Katija, "Fathomnet: An underwater image training database for ocean exploration and discovery," *CoRR*, vol. abs/2007.00114, 2020. [Online]. Available: https://arxiv.org/abs/2007.00114

[20] D. M. Markowitz, R. Laha, B. P. Perone, R. D. Pea, and J. N. Bailenson, "Immersive Virtual Reality field trips facilitate learning about climate change," *Frontiers in Psychology*, vol. 9, no. NOV, 2018.

[21] The Hydrous, "Immerse," 2019. [Online]. Available: https://thehydro.us/vr2019

[22] K. J. Mengerink, C. L. Van Dover, J. Ardron, M. Baker, E. Escobar-Briones, K. Gjerde, J. A. Koslow, E. Ramirez-Llodra, A. Lara-Lopez, D. Squires, T. Sutton, A. K. Sweetman, and L. A. Levin, "A Call for Deep-Ocean Stewardship," *Science*, vol. 344, no. 6185, pp. 696 LP – 698, may 2014. [Online]. Available: http://science.sciencemag.org/content/344/6185/696.abstract